

Factors associated with seroprevalence of SARS-CoV-2 in Orange County

For simplicity let

O_i be the odds of testing positive for COVID-19 in Orange County.

$$\vec{\beta}_{\text{Age Group}} = (\beta_{\text{Age}[18-24]}, \beta_{\text{Age}[25-29]}, \beta_{\text{Age}[30-34]}, \beta_{\text{Age}[35-39]}, \beta_{\text{Age}[40-44]}, \beta_{\text{Age}[45-49]}, \beta_{\text{Age}[50-54]}, \beta_{\text{Age}[55-59]}, \beta_{\text{Age}[60-64]}, \beta_{\text{Age}[65-69]}, \beta_{\text{Age}[70-74]}, \beta_{\text{Age}[75-79]}, \beta_{\text{Age}[80-84]}, \beta_{\text{Age}[85+]})$$

$$\vec{\beta}_{\text{Race/ethnicity}} = (\beta_{\text{Asian}}, \beta_{\text{Black}}, \beta_{\text{Hispanic}}, \beta_{\text{PacificIslander}}, \beta_{\text{Unknown}})$$

$$\vec{\beta}_{\text{College}} = (\beta_{\% \text{ with College Degree Quartile 2}}, \beta_{\% \text{ with College Degree Quartile 3}}, \beta_{\% \text{ with College Degree Quartile 4}})$$

$$\vec{\beta}_{\text{Insurance}} = (\beta_{\% \text{ with Insurance Quartile 2}}, \beta_{\% \text{ with Insurance Quartile 3}}, \beta_{\% \text{ with Insurance Quartile 4}})$$

Model 0:

$$\begin{aligned} \log(O_i) = & \beta_0 + \vec{\beta}_{\text{Age Group}} \overrightarrow{\text{Age Group}}_i + \beta_{\text{Gender}} \text{Gender}_i + \vec{\beta}_{\text{Race/ethnicity}} \overrightarrow{\text{Race/ethnicity}}_i \\ & + \vec{\beta}_{\text{College}} \overrightarrow{\% \text{ with College Degree Quartile}}_i \\ & + \vec{\beta}_{\text{Insurance}} \overrightarrow{\% \text{ with Medical Insurance Quartile}}_i \\ & + \beta_{\text{Population Density}} \text{Population Density}_i + \beta_{\text{House Crowding}} \text{House Crowding}_i + \beta_{\text{Median Income}} \text{Median Income}_i, \end{aligned} \quad (1)$$

without a random intercept for zip code.

Model 1:

$$\begin{aligned} \log(O_i) = & \beta_0 + \vec{\beta}_{\text{Age Group}} \overrightarrow{\text{Age Group}}_i + \beta_{\text{Gender}} \text{Gender}_i + \vec{\beta}_{\text{Race/ethnicity}} \overrightarrow{\text{Race/ethnicity}}_i \\ & + \vec{\beta}_{\text{College}} \overrightarrow{\% \text{ with College Degree Quartile}}_i \\ & + \vec{\beta}_{\text{Insurance}} \overrightarrow{\% \text{ with Medical Insurance Quartile}}_i \\ & + \beta_{\text{Population Density}} \text{Population Density}_i + \beta_{\text{House Crowding}} \text{House Crowding}_i + \beta_{\text{Median Income}} \text{Median Income}_i, \end{aligned} \quad (2)$$

with a random intercept for zip code.

Model 2:

$$\begin{aligned} \log(O_i) = & \beta_0 + \vec{\beta}_{\text{Age Group}} \overrightarrow{\text{Age Group}}_i + \beta_{\text{Gender}} \text{Gender}_i + \vec{\beta}_{\text{Race/ethnicity}} \overrightarrow{\text{Race/ethnicity}}_i \\ & + \vec{\beta}_{\text{College}} \overrightarrow{\% \text{ with College Degree Quartile}}_i \\ & + \vec{\beta}_{\text{Insurance}} \overrightarrow{\% \text{ with Medical Insurance Quartile}}_i \\ & + \beta_{\text{Population Density}} \text{Population Density}_i + \beta_{\text{House Crowding}} \text{House Crowding}_i + \beta_{\text{Median Income}} \text{Median Income}_i \\ & + \beta_{\% \text{ of Zip Code SARS-CoV-2+}} \% \text{ of Zip Code SARS-CoV-2+}_i, \end{aligned} \quad (3)$$

without a random intercept for zip code.

Model 3:

$$\log(O_i) = \beta_0 + \beta_{\% \text{ of Zip Code SARS-CoV-2+}} \% \text{ of Zip Code SARS-CoV-2+}_i, \quad (4)$$

without a random intercept for zip code.

Model 4:

$$\log(O_i) = \beta_0 + \overrightarrow{\beta}_{\text{Age Group}} \overrightarrow{\text{Age Group}}_i + \beta_{\text{Gender}} \text{Gender}_i + \overrightarrow{\beta}_{\text{Race/ethnicity}} \overrightarrow{\text{Race/ethnicity}}_i + \beta_{\% \text{ of Zip Code SARS-CoV-2}} \% \text{ of Zip Code SARS-CoV-2}_i, \quad (5)$$

without a random intercept for zip code.

Table 1: Model comparison using BIC shows negligible difference in modeling odds of testing sero-positive for COVID-19 in Orange County. Therefore the simpler model, Model 0, was chosen.

	Degrees of Freedom	BIC
Model 0	24	2277.576
Model 1	25	2285.567
Model 2	25	2284.740
Model 3	2	2131.675
Model 4	16	2220.823

Table 2: Model 0 regression estimation of adjusted odds ratio of testing sero-positive for SARS-CoV-2 in Orange County.

	Counts		Adjusted Odds Ratio [*] with (95% CI [†])
	SARS-CoV-2+	Total	
Age			
18-24	19 (5.43%)	158 (5.35%)	Reference
25-29	31 (8.86%)	234 (7.92%)	1.085 (0.58, 2.04)
30-34	33 (9.43%)	275 (9.31%)	0.97 (0.52, 1.81)
35-39	35 (10%)	328 (11.1%)	0.848 (0.46, 1.57)
40-49	83 (23.71%)	651 (22.04%)	1.069 (0.62, 1.86)
50-59	82 (23.43%)	659 (22.31%)	1.083 (0.62, 1.89)
60-69	46 (13.14%)	418 (14.15%)	1.019 (0.56, 1.86)
70-79	18 (5.14%)	188 (6.36%)	0.925 (0.45, 1.9)
80+	3 (0.86%)	43 (1.46%)	0.632 (0.17, 2.3)
Gender			
Female	222 (63.43%)	1668 (56.47%)	Reference
Male	128 (36.57%)	1286 (43.53%)	0.748 (0.59, 0.95)
Race/ethnicity [‡]			
White	108 (30.86%)	1228 (41.57%)	Reference
Asian	47 (13.43%)	435 (14.73%)	1.231 (0.84, 1.8)
Black	5 (1.43%)	42 (1.42%)	1.281 (0.49, 3.38)
Hispanic	162 (46.29%)	1010 (34.19%)	1.535 (1.16, 2.02)
Pacific Islander	3 (0.86%)	12 (0.41%)	3.938 (1.04, 14.97)
Unknown	25 (7.14%)	227 (7.68%)	1.244 (0.78, 1.99)
% with College Degree [§]			
1st Quartile	158 (45.14%)	937 (31.72%)	Reference
2nd Quartile	92 (26.29%)	893 (30.23%)	0.945 (0.63, 1.42)
3rd Quartile	59 (16.86%)	644 (21.8%)	1.091 (0.62, 1.94)
4th Quartile	41 (11.71%)	480 (16.25%)	1.069 (0.55, 2.06)
% with Insurance			
1st Quartile	154 (44%)	928 (31.42%)	Reference
2nd Quartile	91 (26%)	812 (27.49%)	0.931 (0.65, 1.34)
3rd Quartile	54 (15.43%)	597 (20.21%)	0.909 (0.52, 1.58)
4th Quartile	51 (14.57%)	617 (20.89%)	0.853 (0.48, 1.53)
Population Density (1000ppl/km ²)			0.998 (0.79, 1.26)
House Crowding			1.013 (0.99, 1.04)
Median Income (std. dev.)			0.79 (0.61, 1.03)

^{*} Model intercept represents odds of testing sero-positive for SARS-CoV-2 for a white female diagnosed with SARS-CoV-2 in the 18-24 age group in a zip code in the first quartile of college degree and insured with the average population density in Orange County. The odds of this individual testing sero-positive is estimated to be 0.102 (0.046,0.222)

[†] 95% confidence interval computed with robust standard errors

[‡] Native American/Native Alaskan race group not included in analysis due to lack of data, no individual of this race group tested seropositive.

[§] The esimated percent of people with a bachelor's degree, and similarly the estimated percent of people with medical insurance, in an individual's zip code

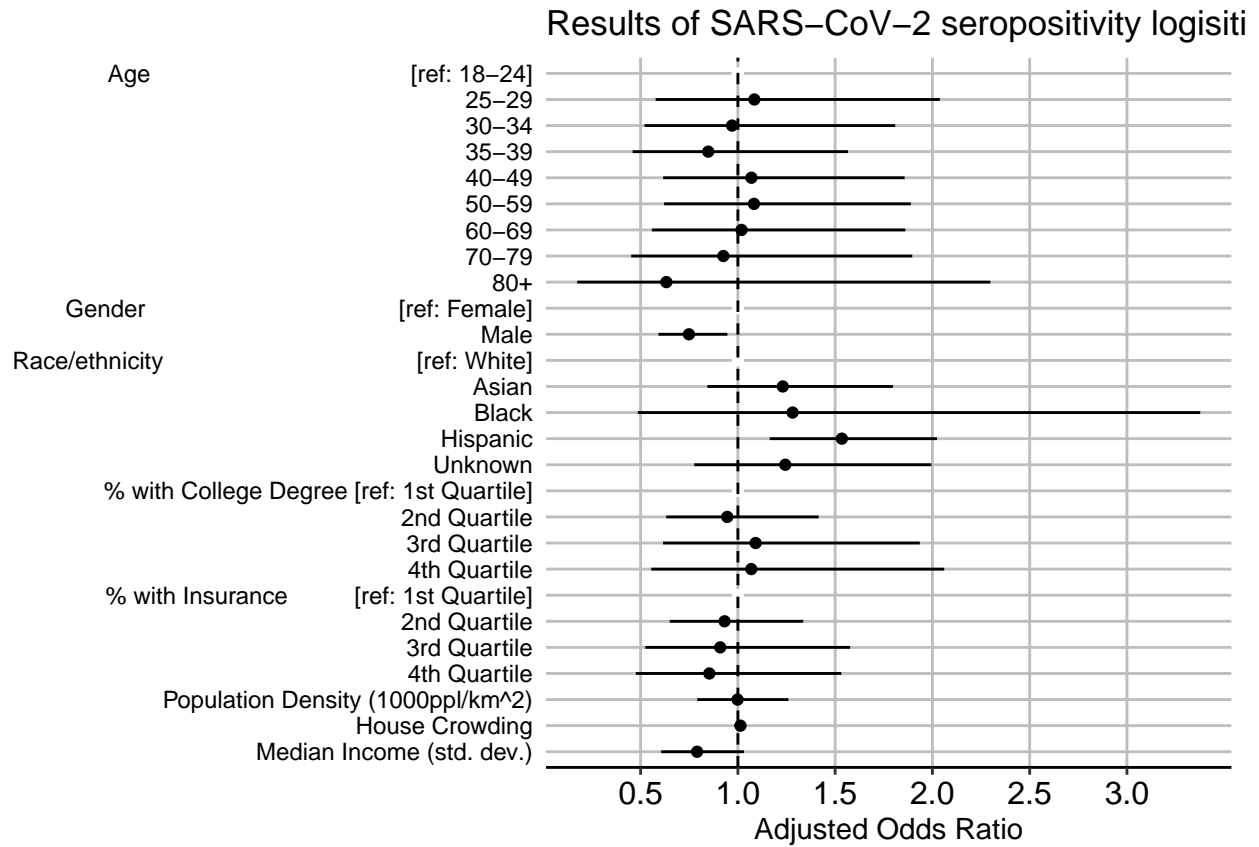


Figure 1: Logistic model 0 results for odds of testing sero-positive for SARS-CoV-2 in Orange County. Pasific Islander adjusted odds ratio omitted due to width of confidence interval (1.05, 15.14).

Table 3: Model 2 regression estimation of adjusted odds ratio of testing sero-positive for SARS-CoV-2 in Orange County.

	Counts		Adjusted Odds Ratio [*] with (95% CI [†])
	SARS-CoV-2+	Total	
Age			
18-24	19 (5.43%)	158 (5.35%)	Reference
25-29	31 (8.86%)	234 (7.92%)	1.085 (0.58, 2.04)
30-34	33 (9.43%)	275 (9.31%)	0.973 (0.52, 1.81)
35-39	35 (10%)	328 (11.1%)	0.847 (0.46, 1.56)
40-49	83 (23.71%)	651 (22.04%)	1.075 (0.62, 1.87)
50-59	82 (23.43%)	659 (22.31%)	1.085 (0.62, 1.89)
60-69	46 (13.14%)	418 (14.15%)	1.018 (0.56, 1.86)
70-79	18 (5.14%)	188 (6.36%)	0.932 (0.45, 1.91)
80+	3 (0.86%)	43 (1.46%)	0.641 (0.18, 2.32)
Gender			
Female	222 (63.43%)	1668 (56.47%)	Reference
Male	128 (36.57%)	1286 (43.53%)	0.746 (0.59, 0.94)
Race/ethnicity [‡]			
White	108 (30.86%)	1228 (41.57%)	Reference
Asian	47 (13.43%)	435 (14.73%)	1.245 (0.85, 1.82)
Black	5 (1.43%)	42 (1.42%)	1.278 (0.48, 3.37)
Hispanic	162 (46.29%)	1010 (34.19%)	1.538 (1.17, 2.03)
Pacific Islander	3 (0.86%)	12 (0.41%)	3.893 (1.04, 14.65)
Unknown	25 (7.14%)	227 (7.68%)	1.246 (0.78, 2)
% with College Degree [§]			
1st Quartile	158 (45.14%)	937 (31.72%)	Reference
2nd Quartile	92 (26.29%)	893 (30.23%)	0.975 (0.65, 1.46)
3rd Quartile	59 (16.86%)	644 (21.8%)	1.146 (0.64, 2.04)
4th Quartile	41 (11.71%)	480 (16.25%)	1.146 (0.59, 2.22)
% with Insurance			
1st Quartile	154 (44%)	928 (31.42%)	Reference
2nd Quartile	91 (26%)	812 (27.49%)	0.979 (0.67, 1.43)
3rd Quartile	54 (15.43%)	597 (20.21%)	0.989 (0.56, 1.76)
4th Quartile	51 (14.57%)	617 (20.89%)	0.95 (0.51, 1.76)
Population Density (1000ppl/km ²)			1.022 (0.81, 1.29)
House Crowding			1 (0.96, 1.04)
Median Income (std. dev.)			0.757 (0.57, 1)
% of Zip Code SARS-CoV-2+ [¶]			1.249 (0.8, 1.96)

^{*} Model intercept represents odds of testing sero-positive for SARS-CoV-2 for a white female diagnosed with SARS-CoV-2 in the 18-24 age group in a zip code in the first quartile of college degree and insured with the average population density, and average percent of SARS-CoV-2 positive individuals in Orange County. The odds of this individual testing sero-positive is estimated to be 0.081 (0.034,0.192)

[†] 95% confidence interval computed with robust standard errors

[‡] Native American/Native Alaskan race group not included in analysis due to lack of data, no individual of this race group tested seropositive

[§] The esimated percent of people with a bachelor's degree, and similarly the estimated percent of people with medical insurance, in an individual's zip code

[¶] Number of individuals who tested positive in individual's zip code reported to OC Public Health Department from March 1st to August 16th, divided by estimated population of zipcode

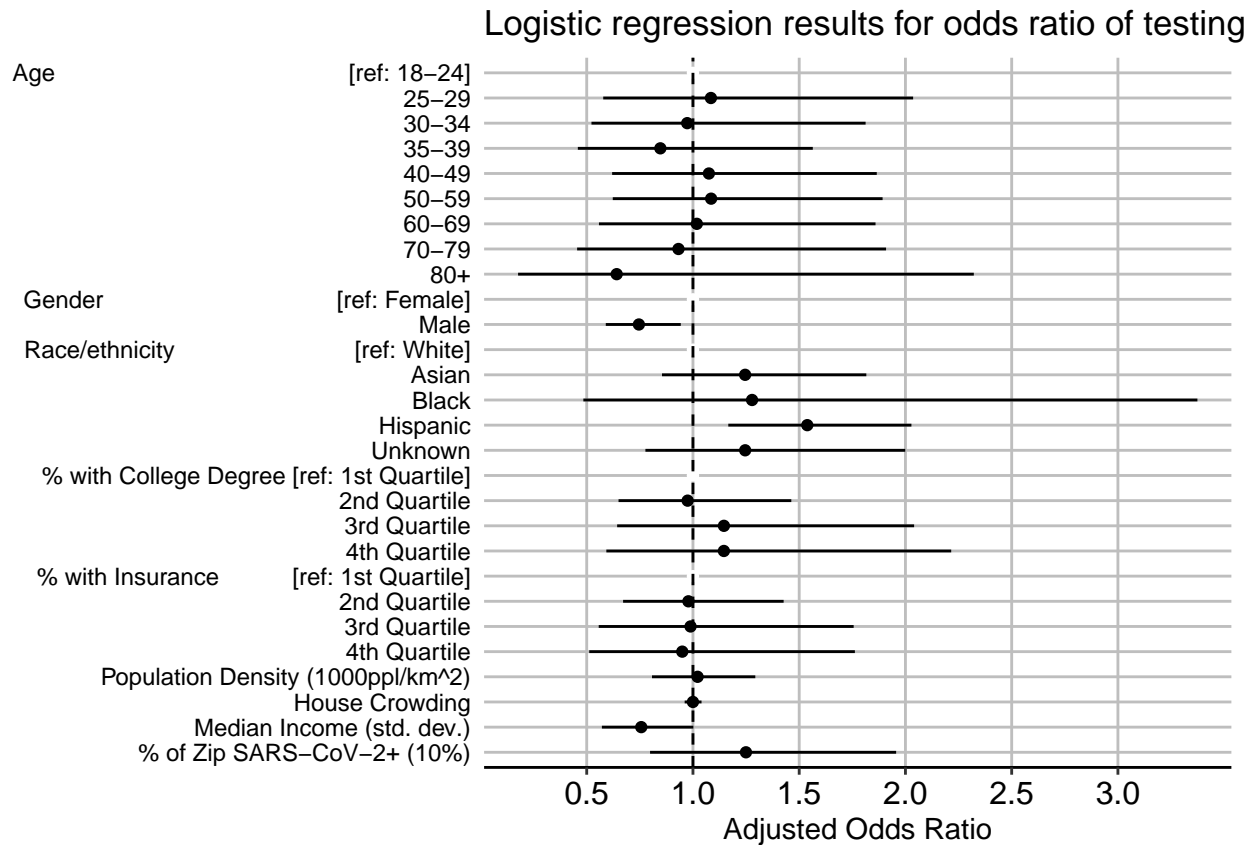


Figure 2: Logistic model 2 results for odds of testing sero-positive for SARS-CoV-2 in Orange County. Pasific Islander adjusted odds ratio omitted due to width of confidence interval (1.05, 15.14).

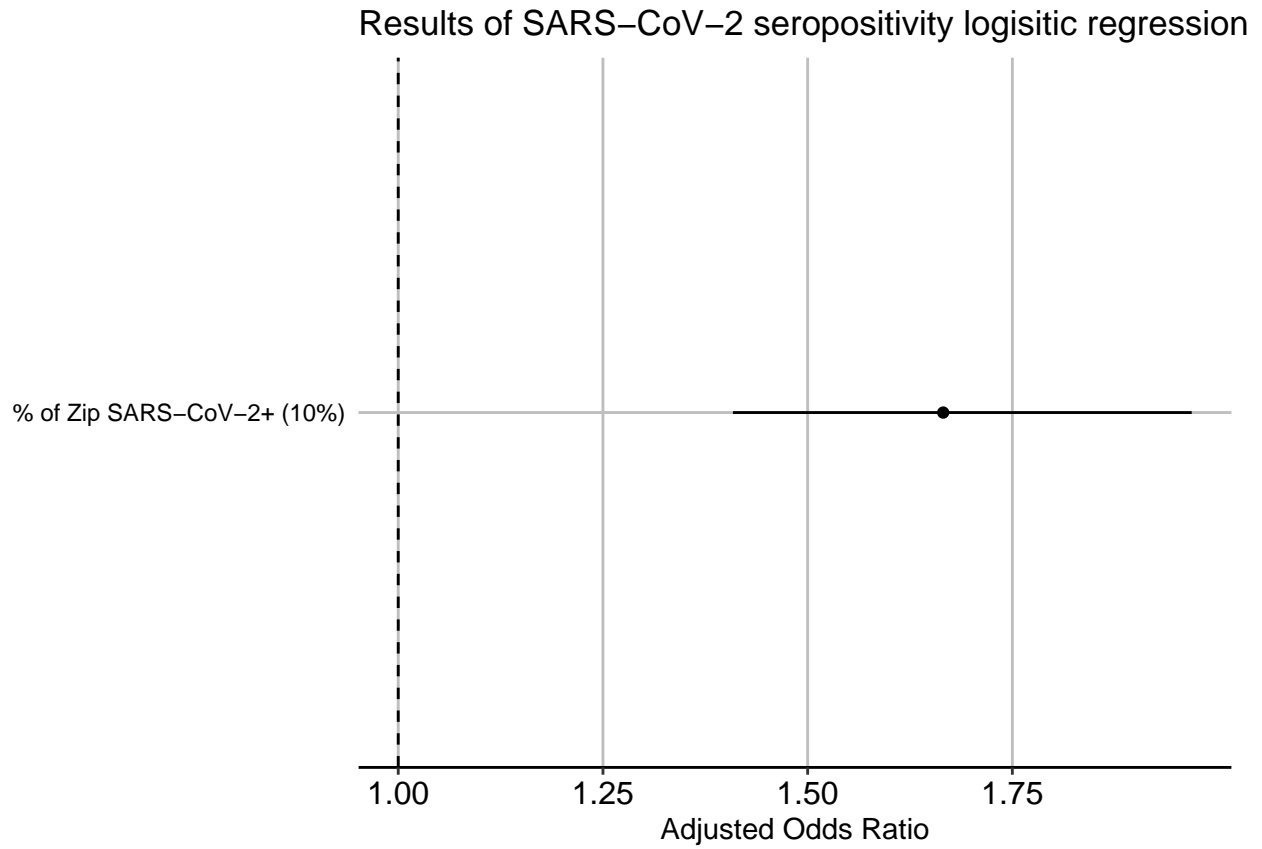


Figure 3: Logistic model 3 results for odds of testing sero-positive for SARS-CoV-2 in Orange County modeled only by percent of their zip code that has tested positive for SARS-CoV-2

Table 4: Model 4 regression estimation of adjusted odds ratio of testing sero-positive for SARS-CoV-2 in Orange County.

	Counts		Adjusted Odds Ratio [*] with (95% CI [†])
	SARS-CoV-2+	Total	
Age			
18-24	19 (5.43%)	158 (5.35%)	Reference
25-29	31 (8.86%)	234 (7.92%)	1.116 (0.6, 2.09)
30-34	33 (9.43%)	275 (9.31%)	0.982 (0.53, 1.82)
35-39	35 (10%)	328 (11.1%)	0.837 (0.46, 1.54)
40-49	83 (23.71%)	651 (22.04%)	1.062 (0.62, 1.83)
50-59	82 (23.43%)	659 (22.31%)	1.069 (0.62, 1.85)
60-69	46 (13.14%)	418 (14.15%)	0.991 (0.55, 1.8)
70-79	18 (5.14%)	188 (6.36%)	0.929 (0.46, 1.9)
80+	3 (0.86%)	43 (1.46%)	0.678 (0.19, 2.42)
Gender			
Female	222 (63.43%)	1668 (56.47%)	Reference
Male	128 (36.57%)	1286 (43.53%)	0.73 (0.58, 0.92)
Race/ethnicity [‡]			
White	108 (30.86%)	1228 (41.57%)	Reference
Asian	47 (13.43%)	435 (14.73%)	1.28 (0.88, 1.86)
Black	5 (1.43%)	42 (1.42%)	1.304 (0.5, 3.41)
Hispanic	162 (46.29%)	1010 (34.19%)	1.59 (1.21, 2.09)
Pacific Islander	3 (0.86%)	12 (0.41%)	3.61 (0.99, 13.11)
Unknown	25 (7.14%)	227 (7.68%)	1.253 (0.78, 2.01)
% of Zip Code SARS-CoV-2+ [§]			1.486 (1.24, 1.78)

* Model intercept represents odds of testing sero-positive for SARS-CoV-2 for a white female diagnosed with SARS-CoV-2 in the 18-24 age group in a zip code in the first quartile of college degree and insured with the average population density, and average percent of SARS-CoV-2 positive individuals in Orange County. The odds of this individual testing sero-positive is estimated to be 0.073 (0.04, 0.131)

[†] 95% confidence interval computed with robust standard errors

[‡] Native American/Native Alaskan race group not included in analysis due to lack of data, no individual of this race group tested seropositive

[§] Number of individuals who tested positive in individual's zip code reported to OC Public Health Department from March 1st to August 16th, divided by estimated population of zipcode

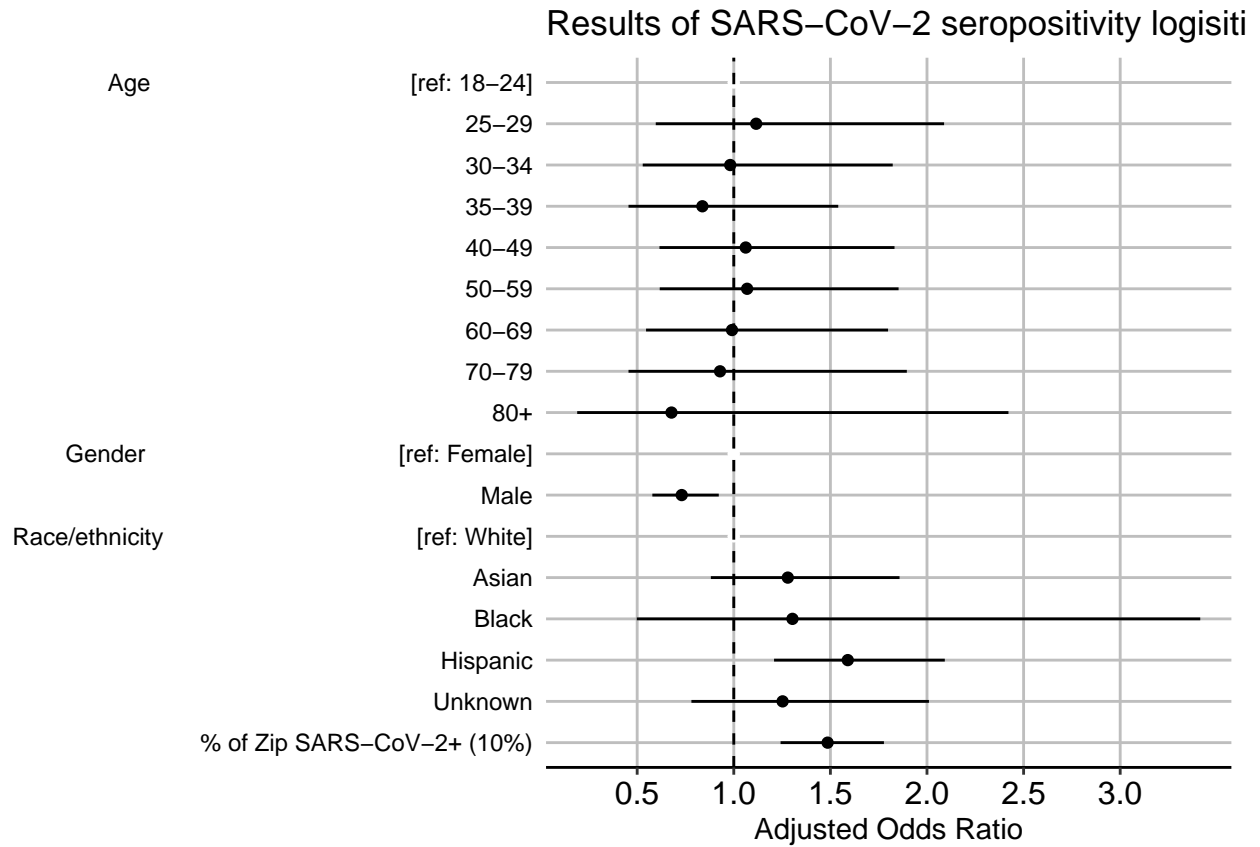


Figure 4: Logistic model 4 results for odds of testing sero-positive for SARS-CoV-2 in Orange County. Pasific Islander adjusted odds ratio omitted due to width of confidence interval (1.05, 15.14).